Algorithms for Optimum Detection of Signals in Gaussian Noise

C.R. Baker

Department of Statistics
University of North Carolina
Chapel Hill, N.C. 27599-3260

LISS 48

March, 1991

ABSTRACT

Algorithms are presented for detection of signals in Gaussian noise. The signals can be Gaussian or nonGaussian. The algorithms are derived from a general solution to the continuous-time problem, and are approximations to the continuous-time likelihood ratio. They do not require knowledge of the probability distributions for the signal-plus-noise process, but instead require knowledge (or estimation) of a function. Independent sampling is not assumed. One algorithm is fully adaptive to the signal-plus-noise process. The algorithms have the potential of providing significant performance improvements, as compared to classical detection methods, when the signal-plus-noise process is broadband (stationary or nonstationary), and particularly when it is nonGaussian.

Research supported by ONR Grant N00014-89-J-1175 and NSF Grant NCR-8713727

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for Public Release: Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: Department of Statistics; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code): University of North Carolina, Chapel Hill, North Carolina 27514	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: Office of Naval Research; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N00014-89-J-1175
8c. ADDRESS (City, State and ZIP Code): Statistics & Probability Program, Arlington, VA 22217	10. SOURCE OF FUNDING NOS. — PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT NO.
11. TITLE (Include Security Classification): Algorithms for Optimum Detection of (cont.)	
12. PERSONAL AUTHOR(S): C.R. Baker	
13a. TYPE OF REPORT: TECHNICAL; 13b. TIME COVERED: FROM _____ TO _____	14. DATE OF REPORT (Yr., Mo., Day): 1991, March; 15. PAGE COUNT: 31
16. SUPPLEMENTARY NOTATION	
17. COSATI CODES — FIELD / GROUP / SUB. GR.	18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number): NonGaussian signal detection, Likelihood ratios, Sonar signal processing

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

TITLE CONT.: Signals in Gaussian Noise

Algorithms are presented for detection of signals in Gaussian noise. The signals can be Gaussian or nonGaussian. The algorithms are derived from a general solution to the continuous-time problem, and are approximations to the continuous-time likelihood ratio. They do not require knowledge of the probability distributions for the signal-plus-noise process, but instead require knowledge (or estimation) of a function. Independent sampling is not assumed. One algorithm is fully adaptive to the signal-plus-noise process. The algorithms have the potential of providing significant performance improvements, as compared to classical detection methods, when the signal-plus-noise is broadband (stationary or nonstationary), and particularly when it is nonGaussian.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☒ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION	
22a. NAME OF RESPONSIBLE INDIVIDUAL: C.R. Baker	22b. TELEPHONE NUMBER (Include Area Code): (919) 962-2189	22c. OFFICE SYMBOL

DD FORM 1473, 83 APR — EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

1. Introduction

Detection of stochastic signals in non-white Gaussian noise can be reasonably characterized as the fundamental problem in sonar signal detection. It also arises in many other applications; for example, in radar problems involving clutter, a Gaussian model is frequently used to model the clutter. The typical problem arises in continuous time with observation over some time interval, although modern signal-processing capabilities usually result in a discrete-time detection algorithm. The desired solution is an algorithm based on the likelihood ratio, which is of course the optimum operation under various criteria. In particular, any strictly monotone increasing function of the likelihood ratio provides a statistical test satisfying the Neyman-Pearson criterion: maximization of probability of detection (P_D) for any fixed value of the probability of false alarm (P_{FA}). Likelihood-ratio solutions have been known for some special cases, as discussed in the next section of this paper. However, these special cases do not apply to a large proportion of the problems encountered in applications.

For discrete-time finite-sample-size data, the likelihood ratio is easy to construct; it is simply the ratio of the (multivariate) S+N probability density function to the multivariate N (noise) probability density function. Of course, this requires not only that both density functions exist, but also (for implementation) that they be known. "Ah, there's the rub;" these density functions are typically not known in sonar.

This paper presents new algorithms for this problem. They are based on recent and very general mathematical solutions to the continuous-time problem, and are obtained in discrete-time form as approximations to the exact continuous-time likelihood ratio. They are in functional form. Their optimality has the unusual property of not requiring any prior knowledge of

the signal-plus-noise (S+N) statistical properties; instead, it is required that one know the drift function of a diffusion. The algorithms can be used for passive or active detection.

The mathematical solution to the problem of discriminating between a Gaussian process (noise) and a second stochastic process of a general type is given in [1], under very mild assumptions on the noise process. The results include conditions for existence of the likelihood ratio, and expressions for the likelihood ratio when it exists. The analysis in [1] is carried out for the continuous-time problem; the results for the likelihood ratio require knowledge that is not generally available in applications. Thus, effective employment of these results in operational systems requires that a discrete-time approximation be developed, and that such an approximation be given in terms of quantities that can reasonably be expected to be available from either prior knowledge or estimation from observed data.

A first approximation to the general solution is given in [2]. However, it has the disadvantage that it requires determination of the eigenvectors and associated eigenvalues of the noise matrix, a problem that can be extremely difficult when using data containing many sample points, and which does not lend itself to a recursive formulation. An implementation was given in [3] which does not suffer these disadvantages. However, that paper did not give an adaptive formulation, nor a formulation based on time-varying properties of the S+N process. Neither [2] nor [3] contains a rigorous justification of the approximations given; in particular, both omit the description of key points that are crucial for the justification of the algorithms. Conversely, the results for the continuous-time problem are stated in mathematical language which is likely to be unfamiliar to many readers interested primarily in applications.

All of these omissions are remedied in the present paper. Two algorithms are presented; one is fully adaptive to the S+N process; the other requires prior knowledge of a diffusion drift function which (after filtering) represents the signal process. Both algorithms require knowledge of the noise covariance matrix, but not of its eigenvalues or eigenvectors. It will be assumed throughout that the noise covariance matrix is known and that the noise has zero mean.

The development here is aimed toward readers who are interested in the actual development of detection algorithms. Given the theoretical basis of the algorithms, the adaptive feature of one, and the many alternatives available for their implementation, they appear to provide a practical solution to this important class of detection problems, with potential performance definitely superior to that of detection methods currently in use when the signal-plus-noise process is braodband. In order that their potential be fully developed, it is important that it be appreciated by those who actually develop surveillance systems. Thus, we have minimized the mathematical jargon so far as possible, have replaced theorems by descriptions and interpretations of results, and have emphasized the applications aspects of the results instead of the mathematical side. However, only so much can be done in this direction; we hope that those strongly interested in the problem will not be deterred by some of the unavoidable complexities. A more mathematical development of the results obtained here, along with additional results, is contained in [22].

In addition to the results described here, we have obtained similar results for detection when the noise is a Gaussian mixture (spherically invariant process); such processes represent, for example, Gaussian processes having covariance known except for a scale factor. Those results will be

published elsewhere; partial results are contained in [3] and [22].

The results given here have already been obtained from us on a privileged basis by some workers in signal detection, and have subsequently been included by at least one for-profit company in successful proposals for Navy funding on development of detection algorithms for this problem. One of the objectives of preparing this paper is to make the algorithms generally available to everyone involved in detection of nonGaussian (or Gaussian) signals in Gaussian noise. We now have substantial extensions of these algorithms, to a wide class of nonGaussian noise processes; the algorithms given here are the basic building blocks for those algorithms.

2. Previous Results on Detection of Stochastic Signals

The problem of detecting a general signal in Gaussian noise has been of acknowledged importance for many years. The major special cases for which a solution has been obtained are listed below. Here, and throughout the paper, by a *likelihood ratio* we mean a test statistic that is a monotone function of the likelihood ratio. The following cases represent those for which a likelihood ratio or generalized likelihood ratio [4], [5] are known.

(1) The signal is a known waveform (for which the classical matched filter is a likelihood ratio), or a waveform known except for a fixed set of parameters (e.g., amplitude, or phase, or time of arrival) whose statistical distribution is known or can be assumed.

(2) A Gaussian signal-plus-noise process.

(3) A Wiener process as the noise.

(4) Independent signal and noise, with the signal paths being comparatively smooth compared to those of the noise.

Each of these problem classes has its drawbacks as a model for applications in sonar. The first category of problems is not sufficiently rich. The second [6], [7] is a class that is often valid, and whose implementation is fairly simple, requiring only quadratic-plus-linear operations on the data. It requires knowledge of the mean vectors and covariance matrix of both the noise (N) and the signal-plus-noise (S+N) processes. These parameters are generally difficult to obtain in advance for the S+N; another important drawback is that the model itself is not valid in many situations; even in cases where it may be valid, the S+N process can vary over time to having strongly nonGaussian properties, and an algorithm limited to the assumption of Gaussian S+N then does not have the flexibility of optimal operation.

Category (3) [8], [9], [10] has a very attractive formulation for the likelihood ratio, which has led to some questionable utilization. In fact, however, it is not suitable as a model for most applications in sonar. First, the sample paths of the Wiener process are far too irregular to give a reasonable model of ocean noise, being non-differentiable at all sample points. Second, the probabilistic properties of the Wiener process are not found in most applications, as they include those of being Markovian and a martingale.

Conditions for existence of the likelihood ratio for problems fitting the model of (4) have been given in [9] and [11]. An expression for the likelihood ratio is included in [11]. However, this model is also not generally practical; first, many important applications do not have independent signal and noise; second, the expression for the likelihood ratio as given in [11] requires that one know the statistical distributions of the signal, which are not likely to be available.

The difficulty of the general problem has led to the use of suboptimum methods; that is, methods based on criteria other than Neyman-Pearson or minimum Bayes risk. Probably the most widely-used such method assumes quadratic-plus-linear operations, with the choice of the operations based on the deflection criterion. This dates back to at least the early 1940's [12], although a general solution to the problem was obtained only in 1969 [13] (see also [14]). Unfortunately, as shown in [13], the optimum operation from the deflection criterion is not a likelihood ratio test statistic even when the S+N process is Gaussian. Nevertheless, the simplicity of the mathematical model and the generality of the solution has led to many papers being written on various modifications of the deflection.

3. Objectives and Summary of Approach

Before moving to the technical content of the paper, we summarize our objectives in this work. We have sought to obtain a detection algorithm, or algorithms, meeting the following criteria.

(1) Likelihood-ratio based.

(2) Information-preserving. The solution should begin with analysis of the original continuous-time problem. The likelihood-ratio for the continuous-time problem should then be approximated as well as possible in forming a discrete-time algorithm. Independent sampling, which typically destroys information, should not be used unless absolutely necessary.

(3) Implementable. An arcane theorem on the form of the likelihood ratio is not our goal; although such theorems can be expected to form part of the process of obtaining a practical algorithm, they must be developed into expressions that can reasonably be expected to be implemented. Thus,

they must not require information that is not reasonably accessible in applications.

(4) Adaptive. A version of the algorithm should be adaptive, particularly to the S+N process.

As will be seen, all four of these goals have been met. Two algorithms will be given; one is fully adaptive to the S+N process; the second requires prior knowledge of the S+N process, but has more flexibility in its implementation when that information is available. Neither algorithm requires the S+N process to have any particular statistical properties. The S+N process can have any statistical properties, with Gaussian or nonGaussian distributions.

In keeping with the above objectives, a general solution to the continuous-time problem was first obtained. This is described in the following section. The likelihood ratio obtained in that work was then approximated to obtain detection algorithms; this is described in the second section below.

The approach taken can be summarized as follows. First, we obtain a set of necessary conditions and a set of sufficient conditions in order that the likelihood ratio exist between some process Y and our noise process N, for the original continuous-time problem. This is a likelihood ratio for probabilities on the space of observations, which we take to be $L_2[0,1]$. The sufficient conditions are satisfied in typical sonar/radar problems, and, assuming these conditions, we obtain an expression for the likelihood ratio. This mathematical expression for the continuous-time problem will not, in general, be implementable; there will not be sufficient information available in applications. A basic problem is that the original representation (see equation (1) below) of the noise process will typically not be known and

cannot be determined in its original generality. Of course, if one has a mathematical model in which the representation is known, then this is not the case. However, here we are assuming the usual situation in practice: no valid mathematical model giving the representation (1) will be availble. Thus, we make an approximation to this representation (justifying the approximation) and then use the necessary conditions for existence of the likelihood ratio to obtain an expression for the likelihood ratio which is based on the S+N process having a representation as a filtered diffusion. This representation is justified. Finally, using this representation, we approximate the continuous-time likelihood ratio to obtain a discrete-time detection algorithm. The estimation of the drift function of the underlying diffusion then becomes the basic problem in implementing the detection algorithm; there are various ways one may approach this problem.

4. Solution of the Continuous-Time Problem

Summary. We assume a mean-square continuous Gaussian process $N = \{N(t), t \text{ in } [0,1]\}$ such that $N(0) = 0$ with probability one. The only other assumption is that N have finite "spectral multiplicity" (also called Cramér-Hida multiplicity; see [15, 16]). Many processes have finite multiplicity, and any m.s. continuous Gaussian process can be approximated by a finite-multiplicity (in fact, multiplicity-one) process. We obtain conditions for existence of the likelihood ratio between N and a second process Y. Y can be thought of as signal-plus-noise. Then, under the assumption that the likelihood ratio exists, we obtain an expression for it. In the process, we also show that Y must have a representation as $Y = S^{*}+N^{*}$, where N^{*} is a noise with the same family of finite-dimensional distributions as N, but with the important

difference that N^* is a function of Y. This is crucial in justifying the final result for the likelihood ratio. Next we show that $Y = \{Y(t), t \in [0,1]\}$ must (if the likelihood ratio exists) have an approximation as a filtered diffusion. The filter and the drift function of the diffusion then characterize the Y (S+N) process.

This section contains a summary of the solution to the continuous-time problem for deciding whether an observed sample function is due only to noise (a Gaussian process of known covariance) or to another process. This second process may be Gaussian or nonGaussian; no assumptions are made on its statistical properties. The mathematical model will first be described; next, we give the solution to the problem of obtaining conditions for existence of the likelihood ratio; finally, expressions for the likelihood ratio are given. We do not aim for precision and generality in this discussion. Instead, we aim to state the results in such a way that they will be accessible for a reader not having background in the esoteric mathematics of stochastic calculus and measures on linear spaces. A precise and complete description of the mathematical problem, a precise theorem-format statement of the results described here, and complete proofs, are contained in [1].

The noise is assumed to be a mean-square continuous Gaussian process. The observation interval is [0,1], although any finite interval can be used. We assume that the noise is "purely nondeterministic" [15], [16]; in essence, that the noise at time zero has the value zero with probability one. Under these assumptions, it is known [15], [16] that the noise has a representation of the following form:

$$N_t = \sum_{i=1}^{M} \int_0^t F_i(t,s)dB_i(s). \qquad (1)$$

For each of the terms appearing in the above sum, the function F_i is a deterministic function. The processes B_i are Gaussian processes that are mutually independent and have independent increments; they can also be taken to be mean-square continuous. The variance of the B_i process, which we denote by β_i, is non-decreasing on [0,1] and so defines an ordinary measure on [0,1] in the usual way: the measure of an interval (a,b) is equal to $\beta_i(b) - \beta_i(a)$. Let $L_2[\beta_i]$ denote the space of all functions on [0,1] that are square-integrable with respect to the measure defined by β_i. Then, for each fixed i and t, $F_i(t,\cdot)$ belongs to $L_2[\beta_i]$. The integrals in (1), expressed in terms of the stochastic processes B_i, are described in many texts; see, for example, [17] or [18]. Essentially, they can be regarded as Stieltjes-type integrals.

The above representation is due to Hida [15] and Cramér [16]. The number M can be infinite; it is the *multiplicity* of the process. One additional assumption to be made here is that this multiplicity is finite. This does not seem to be a serious restriction; it is known that processes of multiplicity one are dense, in a mean-square sense, among all processes having a representation of the above form [19].

The first question that must be considered is that of obtaining conditions for existence of the likelihood ratio. In the case where the signal-plus-noise is Gaussian, it is known [6], [7] that conditions can be given that are necessary and sufficient. However, if a general nonGaussian S+N process is to be admitted, one obtains a set of necessary conditions and a separate set of sufficient conditions; at first glance, similar; on close examination, quite different.

We now give a descriptive summary of these conditions for existence of the likelihood ratio. They will be stated in a rather imprecise mathematical form and all the aspects will not be given; our objective here is to bring out

the main points and to illustrate the importance of some aspects of these conditions. Readers whose primary interest is in development of the algorithms may nevertheless find a careful reading and mulling-over of these conditions to be valuable in understanding how the subsequent detection algorithms follow from the complicated mathematical results of [1]. Many ad hoc detection algorithms have been proposed for problems such as the one described here, and such algorithms are sometimes useful for purposes other than production of papers. However, the algorithms to be given in the next section are not of this genre; they have their genesis in the conditions for existence of the continuous-time likelihood ratio.

The necessary conditions are given in Theorem 2 of [1]. Here, we shall paraphrase that result in descriptive terms. First, let V be any stochastic process independent of N. Let Y be any process defined on [0,1]. We want conditions that are necessary for the likelihood ratio to exist for discrimination between Y and N. Suppose that Y is *adapted* to the combined processes V and N; in essence, this means that the Y process at time t can be obtained as a function of the V and N processes, up to time t. Then, if the likelihood ratio exists, Y must have a representation of the form

$$Y_t = S_t^* + N_t^* \qquad \text{(with probability one)}$$

for all fixed t in [0,1]. The process N^* has the same family of finite-dimensional distributions as the original noise process, and has a Cramér-Hida representation of the same form:

$$N_t^* = \sum_{i=1}^{M} \int_0^t F_i(t,s)\,dB_i^*(s) \qquad (2)$$

where the B_i^* processes each have the same law as the original B_i process, and the F_i functions are those that appear in the representation (1) for the N

process. However, the N^* process has the property (not generally shared by the N process) that it can be written as a functional of the Y process. Moreover, the signal process (S) must have its sample paths (with probability one) belonging to the reproducing kernel Hilbert space (RKHS; see, for example, [17] or [20]) of the noise covariance function, and must also be adapted to the Y process. The RKHS condition can be interpreted as follows: A function f is in the RKHS of N if and only if the likelihood ratio exists between N and N+f. In essence, the RKHS condition means that f is sufficiently smooth (smoother) as compared to the sample functions of the noise.

A very important part of the above statement is that existence of the likelihood ratio implies that the representation of (Y_t) as given above is with respect to a process (N_t^*) which has the same distributions as (N_t) but has the property that it can be represented as a functional of (Y_t), which is not the case for the original (N_t). In the next section, the crucial importance of this fact will be more apparent; we defer a discussion until that point.

In the above necessary condition for existence of the likelihood ratio, we have introduced an independent process V. In sonar applications, this is immediately interpreted. For active sonar, reflection from a target with random reflecting surfaces would be a function of a process (the random reflecting surfaces) that is essentially independent of the scatterers that produce reverberation. Thus, the received signal Y would depend on the noise process (reverberation plus receiver noise)N and on the target reflection process V. In the case of passive detection, the same rationale holds; this is particularly apparent for broadband signal sources. Thus, the initial assumption that (Y_t) be a functional of V and N has clear physical meaning.

We now turn to a summary of the sufficient conditions for existence of

the likelihood ratio as given in Theorem 3 of [1]. As the above discussion shows, we can begin by assuming that (Y_t) is a functional of (N_t) and a process (V_t) that is independent of (N_t). Under this assumption, the likelihood ratio dP_Y/dP_N will exist if Y has a representation as

$$Y_t = S_t + N_t \qquad \text{(with probability one)}$$

for each fixed t in [0,1], where S is a process having all paths (with probability one) in the RKHS of the noise covariance function, and such that S is adapted to the combined V and N processes.

One may note the difference between the necessary condition and the sufficient conditions. In both cases, Y has a representation as the sum of a signal process and a noise process. In the representation for Y under the sufficient conditions, the noise process is the original noise. In the necessary conditions, the noise process (N^*) has the same finite-dimensional distributions as the original noise process (N), but is a functional of the Y process. The S process has similar representations in both sets of conditions; in particular, in both cases it has sample paths belonging (with probability one) to the RKHS of the noise covariance function. However, in the necessary conditions, it is seen that the process (S_t^*) must be adapted to the Y process (which in turn is adapted to the combined V and N processes), while in the sufficient conditions (S_t) must only be adapted to the combined V and N processes. Thus, the properties of the process S^* appearing in the necessary conditions representation $(Y = S^* + N^*)$ are actually stronger than the properties of the process S appearing in the sufficient conditions representation $(Y = S + N)$.

The sufficient conditions for existence of the likelihood ratio are those which one might expect to be satisfied in the usual sonar or radar problem. That is, the problem is usually one containing signal and additive noise; the

signal process will typically be a functional of the noise process and some independent message process, as previously described. Thus, the assumption that the sufficient conditions are satisfied is quite tenable.

The necessary conditions and the sufficient conditions are quite important. The combination of the sufficient conditions and the known physical properties of typical sonar detection problems permit one to determine a likelihood ratio. Given the form of that likelihood ratio, the necessary conditions are then used to justify a functional representation of the S+N process (Y in the above) which leads to a representation of the S+N process as a filtered diffusion. This will be carried out in the following section. However, we first give an expression for the continuous-time likelihood ratio.

Referring to the original representation (1) of the noise process, one shows the existence of a function m defined as follows. The range of m is contained in the space of vector functions of M components, where each component is a function on [0,1] which is continuous and vanishes at zero. The domain of m is in $L_2[0,1]$. m is not defined everywhere in $L_2[0,1]$, but on a dense subset with probability one. Essentially, m can be interpreted as follows. Let Φ be the map which takes the vector process $(B_1,\ldots,B_M)$ into the noise process N (see equation (1)). This map, which has its domain in the space of vector-valued M-component continuous functions vanishing at zero and its range in $L_2[0,1]$, is also defined in a probability one sense. Then $m[\Phi](x) = x$ with probability one (P_W) for x an M-vector of continuous functions vanishing at zero. Thus, roughly, one can view m as being the inverse of the "function" which maps the vector process $(B_1,\ldots,B_M)$ into the process N.

The sufficient conditions described above are reasonable, given the physical nature of the detection problem. We assume they are satisfied. They

show that we can express S+N (or Y, in the above) in the form

$$Y_t = [FdZ]_t \equiv \sum_{i=1}^{M} \int_0^t F_i(t,s)dZ_i(s)$$

where Z is a vector process with M components, each having continuous sample paths, and with i^{th} component having the form

$$Z_i(t) = \int_0^t v_i(s)d\beta_i(s) + B_i(t)$$

where (v_i), $i \leq M$, are stochastic processes, with v_i havings its sample paths (w.p. 1) in $L_2[\beta_i]$. Now, it can be shown that the likelihood ratio must exist between the two vector processes Z and B; write this likelihood ratio as dP_Z/dP_B. This is a likelihood ratio on the space of M-component vector functions where each component is a continuous function on [0,1] that vanishes at zero. Then, the likelihood ratio $dP_{(S+N)}/dP_N$ has the form (when a sample function is observed)

$$\frac{dP_{S+N}}{dP_N}(x) = \frac{dP_Z}{dP_B}(m[x]) \qquad \text{(with probability one } dP_B\text{)}. \qquad (3)$$

We now come to a very important point. The likelihood ratio given above is based on the assumption that the sufficient conditions are satisfied. However, these conditions assume a representation of Y = S + N, where N is the actual noise process. Our eventual algorithm will be derived by first assuming that the noise process N has multiplicity M = 1 in equation (1), and then representing the process Z above (in multiplicity one) as a process of "diffusion type," having the form

$$Z_t = \int_0^t \sigma[s, Z_u,\ u \leq s]ds + W_t \qquad (4)$$

and

$$Y_t = \int_0^t F(t,s)dZ_s \tag{5}$$

where σ is a deterministic function and W is a standard Wiener process. Such a representation for Y always exists [17] if the likelihood ratio dP_Y/dP_N exists. However, (Z_t) depends (in its statistical properties) only on W. In our following analysis, the noise will be obtained from the Wiener process; that is,

$$N_t = \int_0^t F(t,s)dW_s. \tag{6}$$

Thus, if we were to use this representation, then we would have (after filtering to obtain Y = S + N) an S+N process which is completely determined by the noise; obviously an unrealistic and unsatisfactory representation.

However, we can now appeal to the necessary conditions. According to those conditions, we also have the representation of Y given by

$$Y_t = \int_0^t F(t,s)dZ_s^* \tag{7}$$

where F is as above, and Z^* has the representation

$$Z_t^* = \int_0^t \sigma^*(s,Z_u,\ u \leq s)ds + W_t^* \tag{8}$$

where now W^* is also a standard Wiener process, but is a functional of the S+N process, Y. Moreover, since the statistical distributions of W^* and W are the same, they determine the same probability on C[0,1]; similarly Z and Z^* determine the same probability on $L_2[0,1]$. We can therefore assume (7) and (8), rather than (5) and (6). This is a KEY fact in determining the discrete-time detection algorithm. Without this, the representation derived in the next section would be fallacious, since it would ostensibly be based on

a S+N process that depends only on the noise. This is just one of the several subtleties included in the derivation of the detection algorithm which derive from the complicated mathematics of the continuous time problem, and which are in turn so often dismissed as "just mathematics."

In order to implement the likelihood ratio, it is clear that one must know the likelihood ratio of Z (or Z^*) with respect to B, and also the function m. The likelihood ratio of Z with respect to B can be determined from known results, albeit in functional form. Determination of the unknown parameters and the function m is addressed in the next section.

5. Detection Algorithms

Summary. A discrete-time algorithm is obtained by discretizing the approximation to the continuous-time likelihood ratio. This algorithm depends on the covariance matrix of the noise and the drift function of a diffusion. This diffusion, when filtered, gives an approximation to the S+N process. We show that if the drift function is known, then knowledge of the noise covariance matrix enables one to obtain the same PFA as if the (continuous-time) filter were known, while the PD value will approximate the value that would be obtained if the filter were known (the approximation improving as sampling rate increases).

The preceding results are all for the original continuous-time problem. Although they are very general, they suffer from the usual shortcomings of mathematical research devoted to existence of likelihood ratios: they are quite useless as practical detection algorithms unless one is in the rare (one may say nonexistent) situation of having a reliable and complete mathematical model which includes all the parameters needed to implement the continuous-

time likelihood ratio. However, they form a basis for deriving practical discrete-time detection algorithms that are solidly based on the continuous-time likelihood ratio, and, in particular, which satisfy the four criteria mentioned in Section 3:

(1) Likelihood-ratio-based;

(2) Information-preserving;

(3) Implementable; and

(4) Adaptive to the S+N process.

Two algorithms will now be derived. One meets all four of the above criteria. The other meets all except (4); it requires prior knowledge of the S+N process. However, it can model a wider class of S+N processes than the adaptive algorithm, has the capability of being relatively more easily implemented in recursive form, and should give better performance in many situations when the necessary prior information is available.

We now proceed to the development of these two detection algorithms. Version I, fully adaptive to the S+N process, is based on the following additional assumptions:

(A.1) The noise process has multiplicity M=1, and the process $(B_1(t))$ is the standard Wiener process $(W(t))$; thus $N(t) = \int_0^t F(t,s)dW(s)$, where F is a Volterra kernel with $\int_0^1 \int_0^1 F^2(t,s)dsdt < \infty$;

(A.2) The signal-plus-noise process can be represented as $Y(t) = \int_0^t F(t,s)dZ(s)$, where the process (Z_t) is a diffusion with respect to a Wiener process $\hat{W}$ and has time-invariant drift function, so that

$$Z(t) = \int_0^t \theta[Z(s)]ds + \hat{W}(t). \qquad (9)$$

where $P\{\omega: \int_0^1 \theta^2[Z(\omega,s)]ds < \infty\} = 1$.

The second algorithm, Version II, assumes (A.1) above and that the signal-plus-noise process can be represented as $Y(t) = \int_0^t F(t,s)dZ(s)$, where

$$(A.3) \qquad Z(t) = \int_0^t \sigma[s, Z(s)]ds + \hat{W}(t), \tag{10}$$

where again $\hat{W}$ is a Wiener process, and $P\{\omega: \int_0^1 \sigma^2[s, Z(\omega,s)]^2 ds < \infty\} = 1$.

Assumption (A.1) can be justified by noting that all stationary processes and all discrete-time processes are of unit multiplicity, that any Gaussian vector can be represented as the result of a lower-triangular matrix operating on white Gaussian noise, and that unit multiplicity processes are dense (in $L_2[dPdt]$) in the class of m.s. continuous processes [19]. Moreover, one can show that the assumption (A.1) is satisfied whenever the noise process has a representation $N(t) = \int_0^t F(t,s)dB(s)$, where the variance of $(B(t))$ is an absolutely continuous function on $[0,1]$. To see, this we note that an independent-increment Gaussian process B with $B(0) = 0$ (with probability one) has the representation $B(t) = W[h(t)]$, where $EB^2(t) = h(t)$ and $F(0) = 0$. In our setup, h is continuous, non-decreasing, and we can assume that $h(t) > h(s)$ if $t > s$. Thus,

$$\int_0^t F(t,s)dB(s) = \int_0^t F(t,s)dW[h(s)] = \int_0^{h^{-1}(t)} F(t,h^{-1}(u))dW(u)$$

(setting $h(s) = u$)

$$= \int_0^{t'} G(t',u)dW(u)$$

where $t' = h^{-1}(t)$, $G(t',u) = F(t,h^{-1}(u))$. Note that if $u > t'$, then $u > h^{-1}(t)$, so $h(u) > t$; hence $G(t',u) = 0$ if $u > t'$.

To clarify the significance of the assumptions (A.2) and (A.3), we reference well-known material on the representation of processes (Z_t) such

that the likelihood ratio dP_Z/dP_W exists. From Theorem 7.11 of [19], any such (Z_t) must be a process of "diffusion type":

$$Z(\omega,t) = \int_0^t \tau[s,\ Z(\omega,u),\ u \leq s)]ds + \hat{W}(\omega,t) \qquad (11)$$

where $(\hat{W}_t)$ is a Wiener process adapted to $\underline{g}(Z)$, and τ is a function on $[0,1]\times C_0[0,1]$ such that

i) for all s in [0,1] and all functions x in C[0,1], $\tau(s,x)$ does not depend on x(t) for any t > s;

ii) $P\{\omega: \int_0^1 \tau^2[s,Z(\omega,\cdot)]ds < \infty\} = 1$.

From the necessary conditions described in Section 4, it can be seen that assumption (A.3) reduces to the assumption that the function τ is memoryless. Of course, there is a very large class of processes such that this assumption is satisfied. Note that this is emphatically <u>not</u> equivalent to the assumption that the observed signal-plus-noise process is the solution to a stochastic differential equation. To be precise, given that assumption (A.1) is satisfied, the assumption (A.3) states that the signal-plus-noise process can be represented as a filtered diffusion. Note that (A.3) is actually an approximation to (11); it is necessary in order to estimate likelihood ratio parameters from data.

Assumption (A.3) is of course much weaker than (A.2); the latter assumes not only that the process Z defined above is a diffusion, but that the drift function is time-homogeneous: $\tau[s,\ Z(\omega,u),\ u \leq s] = \theta[Z(\omega,s)]$ a.e. $dP(\omega)ds$.

The reduction of the problem to the class of processes satisfying (A.2) is motivated by the goal of developing a likelihood-ratio-based detection algorithm that can operate without any prior knowledge of the signal properties: completely adaptive to the S+N process.

An algorithm will now be described, assuming (A.1) and (A.3), and uniform sampling of the continuous-time waveform. For the detection problem as defined above, applying the results for the likelihood ratio described in Section 4, and known results for the Wiener process (see, e.g., [17] or [20]), the general form (under a mild restriction) of the likelihood ratio on $L_2[0,1]$ is

$$[dP_{S+N}/dP_N](x) = \lim_n \exp\,[\Lambda^n(\delta_n(m[x]))]$$

where $0 = t_0^n < t_1^n < t_2^n < \ldots < t_n^n = T$ is a partition of $[0,T]$ such that $\sup_j |t_{j+1}^n - t_j^n| \to 0$, $\delta_n(x) \equiv (x(t_1^n), x(t_2^n), \ldots, x(t_n^n))$, and

$$\Lambda^n[\delta_n[m(x)]] = \sum_{i=0}^{n-1} \sigma(i, m[x](t_i^n))(m[x](t_{i+1}^n) - m[x](t_i^n)) \qquad (12)$$
$$- (1/2) \sum_{i=0}^{n-1} \sigma^2(i, m[x](t_i^n))(t_{i+1}^n - t_i^n),$$

m is defined as in Section 4, and the limit exists in the norm of $L_1[P_N]$.

It should be noted that this approximation does not arise from sampling the continuous-time observation. The situation is much more complicated; the approximation is obtained by sampling of the continuous-time function m(x), where x is the observation. m(x) is a continuous function vanishing at zero. The difficulty is that m will generally not be known.

The representation of (N(t)) by $N(t) = \int_0^t F(t,s)dW(s)$ yields that $R_N = FF^*$, where F is the integral operator with F(t,s) as its kernel, and F^* is its adjoint. This can be used to provide an approximation to the function m(x) appearing in (12) that does not require calculation of eigenvalues and eigenvectors.

First, notice that $\langle e_j, f_t \rangle = \int_0^T \int_0^t F(s,u)du\, e_j(s)ds =$ $\int_0^t \int_0^T F(s,u)e_j(s)dsdu = [LF^*e_j](t)$, where $[Lf](t) \equiv \int_0^t f(v)dv$. L is a function

from $L_2[0,1]$ into $C[0,1]$ (the continuous functions on $[0,1]$). The expression for m given in [1] can then be written as

$$m[x](t) = \lim_{k\to\infty} [LF^* \sum_1^k \langle e_j, x\rangle R_N^{-1} e_j](t) \qquad (13)$$

$$= \lim_{k\to\infty} [LF^* R_N^{-1} P_k x](t)$$

where $P_k x$ is the projection of the function x on the subspace in $L_2[0,1]$ spanned by $\{e_1,\ldots,e_k\}$. Since $R_N^{-1} = F^{*-1}F^{-1}$, the preceding becomes $m[x](t) = \lim_{k\to\infty} [LF^{-1}P_k x](t)$.

For almost all sample functions x (either from noise or signal-plus-noise), $m[x](\cdot)$ is a continuous function on $[0,1]$. Thus the map m is a linear operator from $L_2[0,1]$ into $C[0,1]$ whose domain includes (with probability one) all sample functions of the noise and signal-plus-noise processes.

The difficulty in implementation of the approximate likelihood ratio (12) will lie in determining the function σ and linear operator m. σ is a parameter of the signal-plus-noise process, and its estimation is a problem of considerable interest in stochastic processes (as the drift of a diffusion) and in stochastic filtering. The possibly unbounded linear operator m, mapping $L_2[0,1]$ into $C[0,1]$, depends only on the covariance function of the noise. If σ is known or estimated, and if the noise covariance function is known or estimated, then the preceding expressions can be used to obtain a discrete-time finite-sample approximation to the likelihood ratio. We now develop such approximations.

Let $\underline{R}_N^n$ denote the covariance matrix of the noise; one can write $\underline{R}_N^n = \Delta^n \underline{F}_n \underline{F}_n^*$, where the matrix $\underline{F}_n$ is lower triangular and Δ^n is the sampling

interval (uniform sampling). The expression for m given above is of the form

$$m[x](t) = \lim_{k\to\infty} [LF^{-1}P_k x](t),$$

where $R_N = FF^*$, L is the integration operator, and P_k is the projection of x onto the subspace spanned by $\{e_1,\ldots,e_k\}$, where $\{e_n,\ n\geq 1\}$ are o.n. eigenvectors of R_N. Thus, a reasonable procedure is simply to replace this expression by $\underline{m}^n[\underline{x}^n] = \underline{L}_n\underline{F}_n^{-1}\underline{x}^n$, where $\underline{x}^n$ is the observed data vector, an element of $\mathbb{R}^n$, and $\underline{L}_n$ is the summation operator in $\mathbb{R}^n$; $(\underline{L}_n\underline{x}^n)_j = \Sigma_{i=1}^{j} x_i$.

However, this is so far simply an *ad hoc* assumption. Thus, we now examine the relation of $\underline{L}_n\underline{F}_n^{-1}\underline{x}^n$ to m(x), where $\underline{x}^n = (x(\Delta^n), x(2\Delta^n), \ldots, x(n\Delta^n))$.

As discussed in Section 4, $(m\circ\Phi_1)(x) = x$ a.e. $dP_W(x)$, where Φ maps W into N. The distribution of m(N) is accordingly given by P_W, so that the vector $\delta_n[m(N)] = (m(N)(\Delta^n), m(N)(2\Delta^n), \ldots, m(N)(n\Delta^n))$ has probability distribution $P_W\circ\delta_n^{-1}$, where δ_n: $C[0,1] \to \mathbb{R}^n$ is defined by $\delta_n(x) \equiv (x(\Delta^n), x(2\Delta^n), \ldots, x(n\Delta^n))$. Similarly, defining $\underline{m}^n(\underline{x}^n) = \underline{L}_n\underline{F}_n^{-1}\underline{x}^n$, $\underline{m}^n(\underline{N}^n)$ has probability distribution $P_W\circ\delta_n^{-1}$, from the definition of $\underline{F}_n$. Thus, in the preceding expression (12) for $A^n[\delta_n[m(x)]]$, and setting $t_i^n = i\Delta^n$, one can replace $\delta_n[m(x)]$ by $\underline{m}^n[\underline{x}^n]$; with respect tc P_N, the law of $A^n[\underline{m}^n(\underline{x}^n)]$ will be the same as that of $A^n[\delta_n[m(x)]]$.

Next, suppose that the signal is present. In examining the relation between the law of $A^n[\delta_n[m(x)]]$, as given by (12), and the law of $A^n[\underline{m}^n[\underline{x}^n]]$, obtained by substituting $\underline{m}^n(\underline{x}^n)$ for $\delta_n[m(x)]$, we make the following smoothness assumptions:

(a) $\underline{F}_n(i,j) \cong F(i\Delta^n, j\Delta^n)$ for all $i,j \leq n$;

(b) The law of the random vector in $\mathbb{R}^n$ with (i+1) component given by

$\int_0^{i\Delta^n} \sigma[s,Z(s)]ds + \hat{W}([i+1]\Delta^n)$ is approximately the law of the random vector in $\mathbb{R}^n$ with (i+1) component given by

$\Delta^n \sum_{k=1}^{i} \sigma[k\Delta^n, Z(k\Delta^n)] + \hat{W}([i+1]\Delta^n)$, where

$$Z(t) = \int_0^t \sigma(s,Z(s))ds + \hat{W}(t)$$

as in assumption (A.3).

Assumption (a) is essentially equivalent to assuming that

$$\int_0^{(i\wedge j)\Delta^n} F(i\Delta^n,s)F(j\Delta^n,s)ds \cong \Delta^n \sum_{k=1}^{i\wedge j} F(i\Delta^n, k\Delta^n)F(j\Delta^n, k\Delta^n)$$

for every $i,j \leq n$. It is thus an assumption on the smoothness of $\{F(t,\cdot),\ t \text{ in } [0,1]\}$. Note, however, that the smoothness requirement applies, for fixed t, only to $F(t,\cdot)$ restricted to $[0,t]$. Similarly, (b) amounts to a smoothness assumption on σ.

When signal is present, letting $\underline{y}^n$ be the vector such that $y^n(k) = \int_0^{k\Delta^n} F(k\Delta^n,s)dZ(s)$, one has

$$\underline{m}^n[\underline{x}^n] = \underline{L}_n \underline{F}_n^{-1} \underline{y}^n.$$

Assumptions (a) and (b) then give that the law of $\underline{m}^n[\underline{x}^n]$ is approximately that of $\delta_n[m(x)]$, with the law of $\delta_n[m(x)]$ being approximately that of the random vector with (i+1) component given by

$$\Delta^n \sum_{k=1}^{i} \sigma[k\Delta^n, Z(k\Delta^n)] + \hat{W}((i+1)\Delta^n).$$

Thus, with the smoothness assumptions (a) and (b), $\Delta^n[\delta_n[m(x)]]$ and $\Delta^n[\underline{m}^n[\underline{x}^n]]$ have approximately the same distribution w.r.t. P_{S+N}. From the nature of assumptions (a) and (b), it can be seen that (if F and σ are sufficiently smooth) the approximations can be expected to become better (for a fixed observation time) as n increases (Δ^n decreases).

We thus have, under the assumption that σ is known, and with the smoothness assumptions on F and σ, an approximation to the discretized continuous-time likelihood ratio. The probability of false alarm (P_{FA}) calculated under this approximation will be exactly that which one would obtain with a discretized version of the exact continuous-time likelihood ratio (under assumptions A.1 and A.3). Thus, the false alarm rate is exact for a given sampling rate. The probability of detection (P_D) will be an approximation to that which would be obtained using a discretized version of the exact continuous-time likelihood ratio when A.1 and A.3 are satisfied.

In most applications, of course, σ will not be known. We now describe two approaches to obtaining an estimate of σ, corresponding to the two assumptions (A.2) and (A.3), both based on replacing $\delta_n[m[x]]$ with $\underline{m}^n[\underline{x}^n]$ in the expression (12) for the discretized log-likelihood ratio.

First, suppose assumption (A.3) is satisfied, and that a training ensemble of representative signal-plus-noise data is available, consisting of K vectors $\{\underline{x}^i,\ i \le K\}$, each having n components, with $\underline{x}^i(j) = x^i(j\Delta)$. It is assumed that the vectors represent independent samples. One first applies the matrix $\underline{F}_n^{-1}$ to each element of this ensemble. Under the assumptions (a) and (b) above, this gives the ensemble of vectors $\{\underline{\delta Z}^i,\ i \le K\}$, where

$$(\delta Z^i)(j) = \underline{Z}^i[(j+1)\Delta^n] - \underline{Z}^i[j\Delta^n]$$

$$= \Delta^n \sigma[j,\ Z^i(j\Delta^n)] + \hat{W}^i[(j+1)\Delta^n] - \hat{W}^i[j\Delta^n].$$

One now fixes j, and uses the K values having the above expression ($i \leq K$) to estimate $\sigma(j,\cdot)$. Various procedures can be used to carry out the estimation; note that defining $(\delta\hat{W}^i)(j) \equiv \hat{W}^i[(j+1)\Delta^n] - \hat{W}^i[j\Delta^n]$, the set $\{(\delta\hat{W}^i)(j),\ i \leq K\}$ consists of i.i.d., random variables, with each Gaussian, mean zero, variance Δ^n.

With this estimated σ inserted into the expression (12) for A^n, a sample vector $\underline{x}^n$ is observed. $\underline{m}^n[\underline{x}^n]$ is then formed, and used in the expression for A^n to form the test statistic $A^n[\underline{m}^n[x^n]]$.

If a representative training ensemble of signal-plus-noise data is available from which one can estimate σ, or if σ is known from a mathematical model, then the above procedure gives the preferred mechanization (Version II algorithm). In the case of non-stationary signal and noise, obtaining an ensemble of S+N data, properly aligned, can be expected to be difficult. However, if the signal-plus-noise is a stationary process, then one may opt to use a long segment of S+N data to estimate a time-invariant σ; this segment could then be much longer than the observation time over which the detection algorithm is to perform. It can be shown [3] that use of a time-varying σ gives an exact likelihood ratio for the discrete-time problem if the signal-plus-noise process is Gaussian.

Suppose next that nothing is known about the statistical properties of the signal-plus-noise process, and that an ensemble of training data is not available. Version I of the algorithm (12) is then implemented as follows, for a fixed value of n: The observed sample vector $\underline{x}^n$ is first used to estimate a time-homogenous σ; the estimated σ is inserted into the expression (12) for A^n, and then $A^n[\underline{m}^n(\underline{x}^n)]$ is evaluated. This version of the algorithm corresponds to assumption (A.2).

The estimation of σ is made from the single observed sample vector $\underline{x}^n$ which is to be tested for the presence of a signal. One applies $\underline{F}_n^{-1}$ to $\underline{x}^n$; assuming that $\underline{x}^n$ represents signal-plus-noise, this yields a vector $\underline{\delta Z}$, which has the representation (under assumptions (a) and (b))

$$(\delta Z)(j+1) = \Delta^n \sigma[Z(j\Delta^n)] + (\delta\hat{W})(j+1),$$

$$(\delta\hat{W})(j+1) = \hat{W}[(j+1)\Delta^n] - \hat{W}[j\Delta^n].$$

The elements of $\{\delta\hat{W}(j),\ j \leq K\}$ are i.i.d. random variables, with each Gaussian, mean zero, and variance Δ^n.

The preceding discussion will now be summarized. First, σ is either known or else is estimated by one of the two procedures described above when the observation is an n-component vector $\underline{x}^n$. The test statistic, an approximation to the continuous-time log-likelihood ratio under the assumption that the noise has multiplicity $M = 1$, then has the expression (with the definition of Λ^n slightly changed from (12))

$$\Lambda^n(\underline{x}^n) = \sum_{j=0}^{n-1} \sigma(j,\ [(\underline{L}_n \underline{F}_n^{-1} \underline{x}^n)_j])[(\underline{L}_n \underline{F}_n^{-1} \underline{x}^n)_{j+1} - (\underline{L}_n \underline{F}_n^{-1} \underline{x}^n)_j]$$

$$- (\Delta^n/2) \sum_{j=0}^{n-1} \sigma^2(j,\ [(\underline{L}_n \underline{F}_n^{-1} \underline{x}^n)_j] \qquad (14)$$

$$= \sum_{j=0}^{n-1} \sigma(j,\ [(\underline{L}_n \underline{F}_n^{-1} \underline{x}^n)_j])[(\underline{F}_n^{-1} \underline{x}^n)_{j+1}] - (\Delta^n/2) \sum_{j=0}^{n-1} \sigma^2(j,\ [(\underline{L}_n \underline{F}_n^{-1} \underline{x}^n)_j].$$

If now a new data point x_{n+1} is observed, the approximation has the recursive form

$$\Lambda^{n+1}(\underline{x}^{n+1}) = \Lambda^n(\underline{x}) + \sigma[n,\ (\underline{L}_n \underline{F}_n^{-1} x)_n](\underline{F}_{n+1}^{-1} \underline{x})_{n+1} \qquad (15)$$

$$- (\Delta^n/2)\ \sigma^2[n,\ (\underline{L}\ \underline{F}^{-1} \underline{x})_n].$$

The above procedure requires relatively few additional operations when a new data point is observed. The implementation and calculation of A require the following operations. First, the function σ must be known and programmed or estimated from the observation. Given the value of $A^n(\underline{x}^n)$ and the observation $\underline{x}^n = (x_1,\ldots,x_n)$, one stores $A^n(\underline{x}^n)$, $\underline{x}^n$, $\sigma[n, (\underline{L}_n\underline{F}_n^{-1}\underline{x}^n)_n]$, and $(\underline{L}_n\underline{F}_n^{-1}\underline{x}^n)_n$. When the data point x_{n+1} is received, it is only necessary to use the vector $\underline{x}^{n+1}$ to calculate $(\underline{F}_{n+1}^{-1}\underline{x}^{n+1})_{n+1}$, which is simply cross-correlation of the observation vector $\underline{x}^{n+1}$ with row n+1 of $\underline{F}_{n+1}^{-1}$. This number, say b_{n+1}, is then used to form $A^{n+1}(\underline{x}^{n+1})$,

$$A^{n+1}(\underline{x}^{n+1}) = A^n(\underline{x}^n) + \sigma[n, \sum_{i=1}^{n} b_i]b_{n+1} - (A^n/2)\,\sigma^2[n, \sum_{i=1}^{n} b_i]. \qquad (16)$$

Throughout, we have made the assumption that the noise covariance is known. One then knows $\{\underline{F}_n,\ n > 1\}$, and thus $\{\underline{F}_n^{-1},\ n \geq 1\}$. As mentioned, each new observation of a data point requires only cross-correlation of the observed vector, an element of $\mathbb{R}^{n+1}$, with row n+1 of $\underline{F}_{n+1}^{-1}$. It is not necessary to apply the matrix $\underline{F}_{n+1}^{-1}$ to the data vector.

Implementation of the recursive form of the algorithm is done most conveniently when σ is known, or when a training ensemble of S+N data can be used to estimate σ. If one must estimate σ from the observed data (Version I algorithm), then the recursive formulation given above will need modification. Various approaches can then be used for updating the estimate of σ, depending on the amount of storage available, etc.

A pleasing and important feature of the algorithms in that Version II is a likelihood ratio (exact) when the S+N process is Gaussian (see [3]).

The performance of the algorithm can be expected to depend not only on the properties of the data, but also on the sampling rate and the choice of

the specific estimation procedure for estimating σ. Thus, implementation for a particular application should be preceded by an extensive study featuring both simulated and experimental data. Limited computational evaluations of these algorithms have given excellent results with passive sonar data [21].

6. Concluding Remarks

Many discrete-time finite-sample detection algorithms for problems such as discussed here are obtained from consideration of only the discrete-time (and finite-dimensional) problem. It is obviously preferable, if possible, to develop a discrete-time algorithm based on approximations to the likelihood ratio for the continuous-time problem.

However, likelihood ratios for continuous-time problems involve conditions for existence of the likelihood ratio. Unfortunately, studies on existence of likelihood ratios are frequently considered to be of only mathematical interest. As can be seen from the preceding development, appropriate studies on absolute continuity and likelihood ratios for continuous-time problems can be extremely important in developing practical discrete-time detection algorithms.

References

[1] C.R. Baker and A.F. Gualtierotti. Discrimination with respect to a Gaussian process, *Probabil. Theory Rel. Fields*, 71, 159-182 (1986).

[2] C.R. Baker and A.F. Gualtierotti, Detection of nonGaussian signals in Gaussian noise, *Statistical Signal Processing* (J.G. Smith and E.J. Wegman, eds.), 107-116, Marcel Dekker (1984).

[3] C.R. Baker and A.F. Gualtierotti. Likelihood ratios and signal detection for nonGaussian processes, *Stochastic Processes in Underwater Acoustics*, Lecture Notes in Control and Information Science, 85, 154-180, Springer-Verlag, Berlin (1986).

[4] C.W. Helstrom. *Statistical Theory of Signal Detection*, Pergammon, New York (1960).

[5] H.L. van Trees, *Detection, Estimation, and Linear Modulation Theory, Part I*, Wiley, New York (1968).

[6] C.R. Rao and V.S. Varadarajan. Discrimination of Gaussian processes, *Sankhyā Ser. A*, 25, 303-330 (1963).

[7] M. Rosenblatt, ed. *Proceedings of the Symposium on Time Series Analysis*, Wiley, New York (1963). See Chapter 11 (E. Parzen), Chapter 19 (G. Kallianpur and H. Oodaira), Chapter 20 (W.L. Root), and Chapter 22 (A.M. Yaglom).

[8] T.E. Duncan. Likelihood functions for stochastic signals in white noise, *Inf. Control*, 16, 303-310 (1970).

[9] T.T. Kadota and L.A. Shepp. Conditions for absolute continuity between a certain pair of probability measures, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 16, 250-260 (1970).

[10] T. Kailath and M. Zakai. Absolute continuity and Radon-Nikodym derivatives for certain measures relative to Wiener measure, *Ann. Math. Statist.*, 42, 130-140 (1971).

[11] C.R. Baker. On equivalence of probability measures, *Annals of Probabil.*, 1, 690-698 (1973).

[12] J.L. Lawson and G.E. Uhlenbeck. *Threshold Signals*, McGraw-Hill, New York, pp. 161-163 (1950).

[13] C.R. Baker. On the deflection of a quadratic-linear test statistic, *IEEE Trans. Inform. Theory*, 15, 16-21 (1969).

[14] C.R. Baker. Optimum quadratic detection of a random vector in Gaussian noise, *IEEE Trans. Comm. Tech.*, 14, 802-805 (1966).

[15] H. Cramér. On some classes of non-stationary stochastic processes, *Fourth Berkeley Sympos. Math. Stat. and Probabil.*, 2, 57-77, Univ. Calif. (1961).

[16] T. Hida. Canonical representations of Gaussian processes and their representations, *Mem. Coll. Science*, 33A, 109-115, Univ. Kyoto (1960).

[17] R.S. Liptser and A.N. Shiryayev. *Statistics of Random Processes, I. General Theory*, Springer, Berlin-Heidelberg-New York (1977).

[18] J.L. Doob. *Stochastic Processes*, Wiley, New York (1950).

[19] A. Ephremides, A property of random processes with unit multiplicity, *J. Multivariate Analysis*, 7, 525-534 (1977).

[20] J. Memin. Sur quelques problèmes fondamentaux de la théorie du filtrage, *These (troisieme cycle)*, U.E.R. Mathématiques et Informatique, University of Rennes (1974).

[21] C.R. Baker, M.R. Frey, and C.E. Persons, A computational evaluation of new detection algorithms, *J. of Underwater Acoustics*, 41, 457-476 (1991).

[22] _____ and A.F. Gualtierotti, Likelihood-ratio detection of stochastic signalsx, to appear, *Advances in Statistical Signal Processing*. Vol. 2, *Signal Detection*, H.V. Poor and J.B. Thomas, eds. JAI Press.

AD-A256063 Words/Phrases(4 words max) that match Thesaurus Entries

TEXT	THESAURUS
ASSIMILATION	Assimilation
ASSIMILATIONS	Assimilation
CORRECTION	Corrections
CORRECTIONS	Corrections
CYCLES	Cycles
HEATING	Heating
INTERPOLATION	Interpolation
LABORATORY	Laboratories
MODEL	Models
NAVAL RESEARCH	Naval Research
NAVAL RESEARCH LABORATORY	Naval Research Laboratories
NORTH CAROLINA	North Carolina
NUMERICAL ANALYSIS	Numerical Analysis
PROFILES	Profiles
RAINFALL	Rainfall
SOUNDING	Sounding
TIME	Time
UNIVERSITY	Universities

@1@ AD-A256063 93-02
@20@ u
@23@ Assimilation, Corrections, Cycles, Heating, Interpolation, Laboratories, Models, Naval Research, Naval Research Laboratories, North Carolina, Numerical Analysis, Profiles, Rainfall, Sounding, Time, Universities.
@24@ u
@27@ A method of assimilating 3-hourly sounding data is developed and successfully tested in this study. First, the successive corrections scheme of Bratseth (1986) which converges to optimum interpolation, is applied for the numerical analysis of data collected during the Genesis of Atlantic Lows Experiment (GALE). Next, diabatic forcing is incorporated into a vertical mode initialization scheme to proivde more realistic initial conditions and to shorten the spinup time of the Naval Research Laboratory/North Carolina State University (NRL/NCSU) mesoscale mode. Latent-heating profiles are computed from 'spun-up' model-generated and observed rainfall. Finally, the multivariate, successive correction analysis scheme correction analysis scheme and the diabatic initialization procedure are combined with the NRL/NCSU model to form an intermittent data-assimilation system. Assimilations of the GALE data over a 2 1/2-day period were performed with differing update cycles of 3, 6, and 12 h.
@28@ u